LA-UR-01-4549

TITLE: Evaluation of Land Use/Land Cover Datasets for Urban Watershed Modeling

AUTHOR(S): Steven J. Burian, Michael J. Brown, and Timothy N. McPherson

SUBMITTED TO: IWA 5th International Conference on Diffuse/Nonpoint Pollution & Watershed Management, Milwaukee WI, June 10-15, 2001

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

EVALUATION OF LAND USE/LAND COVER DATASETS FOR URBAN WATERSHED MODELING

S. J. Burian*, M. J. Brown**, and T. N. McPherson**

** Department of Civil Engineering, University of Arkansas, 4190 Bell Engineering Center, Fayetteville, AR 72701 USA, Email: sburian@engr.uark.edu*
*** Energy and Environmental Analysis Group, D-4, Los Alamos National Laboratory, Los Alamos, NM 87545 USA*

ABSTRACT

Land use/land cover (LULC) data are a vital component for nonpoint source pollution modeling. Most watershed hydrology and pollutant loading models use, in some capacity, LULC information to generate runoff and pollutant loading estimates. Simple equation methods predict runoff and pollutant loads using runoff coefficients or pollutant export coefficients that are often correlated to LULC type. Complex models use input variables and parameters to represent watershed characteristics and pollutant buildup and washoff rates as a function of LULC type. Whether using simple or complex models an accurate LULC dataset with an appropriate spatial resolution and level of detail is paramount for reliable predictions. The study presented in this paper compared and evaluated several LULC dataset sources for application in urban environmental modeling. The commonly used USGS LULC datasets have coarser spatial resolution and lower levels of classification than other LULC datasets. In addition, the USGS datasets do not accurately represent the land use in areas that have undergone significant land use change during the past two decades. We performed a watershed modeling analysis of three urban catchments in Los Angeles, California, USA to investigate the relative difference in average annual runoff volumes and total suspended solids (TSS) loads when using the USGS LULC dataset versus using a more detailed and current LULC dataset. When the two LULC datasets were aggregated to the same land use categories, the relative differences in predicted average annual runoff volumes and TSS loads from the three catchments were 8 to 14% and 13 to 40%, respectively. The relative differences did not have a predictable relationship with catchment size.

KEYWORDS

Land cover; land use; nonpoint source pollution; watershed modeling

INTRODUCTION

Land use and land cover (LULC) information is a vital dataset for applications such as urban planning, air quality studies, meteorological modeling, and watershed modeling. Land cover refers to the state or physical appearance of the land surface (e.g., grasslands, forest, bare soil, exposed rock, developed land). Land use refers to the specified purpose of land from a human perspective (e.g., high-density residential, commercial services, row-crop agriculture, managed forest, rangeland). Together land use and land cover information suggest specific characteristics of the land surface (e.g., imperviousness, solar reflectivity, vegetation type, building morphology), which can be incorporated into environmental models as distributed or bulk parameterizations.

Using LULC information in environmental modeling has become much easier in recent years because of the increased availability and accessibility of datasets in digital format, the enhanced quality of the datasets, and the improved ability to manage and visualize geospatial information in geographic information systems (GIS). Furthermore, the reduced cost of collecting and analyzing remotely sensed information has made dataset updates more frequent and cost-effective. Environmental modeling, in general, requires accurate LULC datasets to parameterize the physical system being simulated. For urban hydrology and diffuse pollution models, LULC datasets are critical for describing the distribution of percent impervious area, interception, depression storage, and other common hydrological and water quality parameters throughout a watershed (see Table 1).

Table 1. Common watershed models and basic model parameters related to LULC.

SWMM
Subcatchment percent imperviousness, roughness coefficients, depression storage, water quality parameters (e.g., constant EMC, buildup/washoff, rating curve)
SLAMM
Land use type, stochastic pollutant generation parameters
HSPF
Land segment retention capacity, pollutant accumulation, removal, and washoff
STORM
Runoff coefficient, pollutant buildup and washoff

Currently, one important application of LULC datasets in watershed modeling is total maximum daily load (TMDL) development and the corresponding waste load allocation. LULC datasets are a central component in the U.S. Environmental Protection Agency (EPA) BASINS modeling system (http://www.epa.gov/ostwater/BASINS/), which has been used for TMDL development projects. BASINS core dataset contains the U.S. Geological Survey (USGS) / Environmental Protection Agency (EPA) 1:250,000 scale quadrangle LULC GIRAS datasets. LULC datasets can be used to assign hydrologic and water quality parameters from the literature for use in a version of the Hydrologic Simulation Program — Fortran (HSPF) packaged in BASINS. Unfortunately, the relative ease of application of the BASINS modeling system can mask the limitations and inaccuracies of the underlying datasets (e.g., LULC) and prevent improved datasets from being sought and incorporated into the modeling effort.

Another important application of LULC datasets in watershed modeling is the evaluation of LULC change and the associated impacts on water resources. Harbor (1994) introduced a curve number based modeling system to evaluate the potential long-term hydrological impact of land use change. LULC data are needed to estimate the curve numbers and other hydrological parameters in the model. Bhaduri *et al.* (2000) describe an extension of Harbor s work to include nonpoint source pollution modeling in the long-term assessment of land use change. LULC datasets are again a key component of the modeling system. The identification, through monitoring and modeling, of the strong correlation between land use change and environmental impacts has led to a call for the incorporation of ecological principles in the management of land use (Dale *et al.*, 2000).

Considering urban watershed modeling, the need for LULC datasets with fine spatial resolution and a high level of detail has traditionally been considered relatively unimportant for determining pollutant loads. This assertion was based on the results of the Nationwide Urban Runoff Program (NURP) that suggested there was no significant difference in average pollutant concentration between different urban sites (EPA, 1983). The implication of these results was that a single concentration value for an urban area was sufficient for estimating urban runoff pollutant loads (Schueler, 1987). Stormwater monitoring in the past two decades, however, has shown conclusively that constituent concentrations in stormwater can vary considerably between different urban land uses. Stormwater monitoring in Los Angeles, California, for example, has shown the quality characteristics to be significantly different between seven urban land uses (LADPW, 2000), while in Wisconsin, Bannerman *et al.* (1993) used micro-monitoring samplers to demonstrate the differences in stormwater quality from 12 different types of urban surfaces including roofs, streets, parking lots, lawns, and driveways.

Using a single urban land use in a watershed model might be appropriate for some preliminary estimates of runoff volumes and pollutant loadings from a catchment. Charbeneau and Barrett (1998) suggest that for screening purposes, a constant EMC independent of land use can be used to estimate stormwater loads, but they also found that using EMCs from individual land uses provided a more accurate prediction than using a single EMC. Planning and evaluation of diffuse pollution management plans requires the use of more detailed LULC datasets because diffuse pollution sources and control measures are directly linked to land use. But, there is little available guidance for selecting the appropriate land use classification system to use for urban watershed modeling. Marsalek (1978; referenced in Novotny and Olem, 1994) describes several land use categories that were differentiated by their nonpoint source pollution loading potential. Hill et al. (2000) describe an image analysis technique to rapidly derive land-use coverages for urban watershed analyses. These ideas are promising, but more work needs to be performed to determine the impact different land use classification systems have on urban watershed modeling results. This paper briefly reviews several LULC dataset sources with different levels of detail and the effects of their application to urban diffuse pollution modeling and management planning.

LAND USE/LAND COVER DATA SOURCES

LULC datasets can be derived from a variety of sources of information including satellite imagery, low-altitude aerial photographs, and site canvassing. Deriving LULC datasets from satellite imagery is more cost-effective for large study areas than for smaller study areas. Site canvassing is most applicable for small-scale studies requiring a high degree of accuracy (e.g., zoning/growth planning). LULC dataset spatial resolutions generally vary from tens of square meters for local/regional datasets up to a square kilometer or larger for national or global datasets (e.g., the USGS Global Land Cover Database). The level of detail ranges from less than ten to more than one hundred land classifications.

In the U.S., government agencies are the primary source for LULC datasets used in environmental modeling because the datasets are typically nationally consistent and obtainable free-of-charge or for a nominal fee. The USGS is the federal agency primarily responsible for development, maintenance, and distribution of a nationwide LULC dataset. The traditional USGS LULC dataset is available at scales ranging from 1:250,000 to 1:100,000 depending on location. The land use and cover is classified according to the level II classification scheme described by Anderson et al. (1976) (see Table

2). The level II classification scheme has relatively coarse detail, but permits finer classification to level III or IV. The basic sources of land use compilation data are National Aeronautics and Space Administration (NASA) high-altitude aerial photographs, and National High-Altitude Photography (NHAP) program photographs, usually at scales smaller than l:60,000 (USGS, 1990).

Table 2. Anderson classification system used in USGS LULC datasets.

Level I	Level II
1 Urban or Built-up Land	11 Residential 12 Commercial and Services 13 Industrial 14 Transportation, Communications and Utilities 15 Industrial and Commercial Complexes 16 Mixed Urban or Built-up Land 17 Other Urban or Built-up Land
2 Agricultural Land	21 Cropland and Pasture 22 Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas 23 Confined Feeding Operations 24 Other Agricultural Land
3 Rangeland	31 Herbaceous Rangeland 32 Shrub and Brush Rangeland 33 Mixed Rangeland
4 Forest Land	41 Deciduous Forest Land 42 Evergreen Forest Land 43 Mixed Forest Land
5 Water	51 Streams and Canals 52 Lakes 53 Reservoirs 54 Bays and Estuaries
6 Wetland	61 Forested Wetland 62 Nonforested Wetland
7 Barren Land	71 Dry Salt Flats 72 Beaches 73 Sandy Areas Other than Beaches 74 Bare Exposed Rock 75 Strip Mines, Quarries, and Gravel Pits 76 Transitional Areas 77 Mixed Barren Land
8 Tundra	81 Shrub and Brush Tundra 82 Herbaceous Tundra 83 Bare Ground 84 Wet Tundra 85 Mixed Tundra
9 Perennial Snow or Ice	91 Perennial Snowfields 92 Glaciers

The minimum size of polygons depicting all urban or built-up land (categories 11-17), water (51-54), confined feeding operations (23), other agricultural land (24), strip mines, quarries, and gravel pits (75) and urban transitional areas (76), is 4 ha. All other categories have a minimum polygon size of 16 ha (USGS, 1990). In the urban or built-up land and water categories, the minimum width of a feature to be shown is 200 m. Although the minimum-width consideration precludes the delineation of very narrow and very long 4 ha polygons, triangles or other polygons are acceptable if the base of the triangle or minimum width of the polygon is 200 m in length and if the area of the polygon is 4 ha. Exceptions to this specification are limited access highways (14) and all double line rivers (51) on the 1:250,000-scale base, which shall have a minimum width of 92 m. For categories other than urban or built-up land and water , the 16-ha minimum size for delineation requires a minimum-width polygon of 400 m.

The National Land Cover Dataset (NLCD) is currently being introduced for the eastern U.S., with release of the datasets for the remainder of the country scheduled for the near future (edcwww.cr.usgs.gov/programs/lccp/nationallandcover.html). The NLCD is being produced as part of a cooperative project between the USGS and EPA. The goal of the project is to produce a consistent land cover data layer for the conterminous U.S. based on 30-meter Landsat 5 thematic mapper (TM) data and supplemented by various ancillary data (where available). The classification system used for NLCD is a modified form of the Anderson classification system. Unfortunately, using Landsat TM data is inappropriate to derive many of the Anderson level III classes (e.g., urban), which are best derived using aerial photography. Moreover, some Anderson Level II classes are consolidated into a single NLCD class. Overall, the NLCD is a more up-to-date and accurate dataset than the USGS LULC dataset, but it is not yet available for most of the U.S. and could not be incorporated into the evaluation reported in this paper.

In addition to the USGS, local/regional governments and commercial vendors also distribute LULC datasets. In the past decade, local and regional governments have become increasingly involved in the collection and distribution of LULC datasets because of the widespread use of GIS and the need for LULC GIS data for planning. Commercial vendors have evolved in response to the demand for LULC datasets for numerous applications including planning, environmental impact assessment, scientific studies, engineering design and analysis, and mapping. For watershed modeling the appropriate LULC dataset to use depends on the goal of the analysis, the choice of model, the spatial scale, the desired accuracy, the availability of monitoring data for calibration/verification, and the desired detail of the results. Understanding the sensitivity of urban watershed model results to different LULC dataset sources is an important step in the selection of an appropriate LULC dataset for a particular application. The next section in this paper presents preliminary results from an investigation of the influence of LULC dataset type on urban watershed model results.

WATERSHED MODELING EVALUATION

Three urban catchments in Los Angeles, California were used to evaluate the relative differences in model predictions of average annual runoff volumes and pollutant loads when using two different LULC datasets in an urban runoff model. The three catchments (Centinela, Sepulveda, and Ballona) compose the majority of the Ballona Creek watershed, which drains a large part of Los Angeles. Table 3 shows the catchment areas and overall percent imperviousness. The predominant land use type in each catchment is high-density single family residential.

Table 3. Characteristics of the three Los Angeles study catchments.

	Centinela	**Sepulveda**	**Ballona**
Area (km^2)	23.8	61.4	216.7
% Imperviousness	67	51	65

The EPA Storm Water Management Model (SWMM) (Huber and Dickinson, 1988) was used to simulate the stormwater runoff quantity and quality from the three catchments. SWMM models of the three catchments were developed during a previous project (Burian et al., 2000) and only minor modifications related to LULC input data had to be performed during this study. Total suspended solids (TSS) is the quality constituent simulated. The TSS concentration in the runoff is determined by assigning a constant concentration to the runoff from each simulated land use type in the catchment. The Los Angeles County Department of Public Works (LADPW) has collected land-use specific event mean concentrations (EMCs) in the vicinity of the catchments for the past six years. The eight land uses monitored by the LADPW include commercial, high-density single family residential, transportation, light industrial, urban vacant, educational, multifamily residential, and mixed residential. The median EMCs for each land use from the six years of monitoring were used in the SWMM models to simulate the constant concentration from each land use. Since the median EMCs are representative of long-term conditions the models will only provide accurate predictions over the long-term (e.g., average annual conditions) and predictions of pollutant concentrations at the storm event level may not be accurate. The model predictions have been shown to provide reasonable estimates of annual runoff volumes and pollutant loadings when compared to the Ballona Creek stream gauge records (see Burian et al., 2000).

The USGS LULC dataset and a LULC dataset obtained from the Southern California Association of Governments (SCAG) were used to parameterize the three models. The USGS LULC is classified according to the Anderson level II classification system, while the SCAG dataset is classified with more land use categories according to a modified Anderson level III/IV classification system. The SCAG dataset is based on 1993 aerial photographs, has a spatial resolution of _ ha, and has a higher degree of ground truthing than the USGS dataset. Both LULC datasets were aggregated up to seven land use classes corresponding the Anderson level II classification to make them consistent for the simulations. Seven land classes were chosen for this study because of the limited number of classes in the USGS dataset and more importantly the limited number of land use types with reliable stormwater quality information. Table 4 compares the areas of each aggregated land use type for the three catchments for both the USGS and SCAG datasets. For all three catchments the areas of USGS and SCAG land use types are similar except for the urban vacant and the other urban land classes. This is significant because these two land use types have different runoff and pollutant loading characteristics. The similarities between the two datasets suggest the land use in the three catchments has not changed significantly in the past two decades. The differences in area for the urban vacant and other urban are likely due to differences in classification. The land use types monitored by the LADPW do not correspond one-to-one with the seven aggregated land use classes for the modeling. The mapping between the simulated land use types and the types monitored by the LADPW are shown in Table 5.

Table 4. Area (in hectares) of each aggregated land use class in the three catchments.

	Centinela		**Sepulveda**		**Ballona**	
	USGS	**SCAG**	**USGS**	**SCAG**	**USGS**	**SCAG**
Commercial	318	354	808	797	4393	3684
Residential	1396	1387	3033	2889	14233	14004
Transportation	86	101	257	224	281	396
Industrial	301	248	191	132	713	892
Urban Vacant	6	281	1270	1987	1293	2476
Other Urban	272	8	500	33	722	104
Water	0	0	66	65	34	113

For each catchment, thirty years of rainfall from the Los Angeles International Airport (LAX) were used to generate thirty years of runoff. The thirty-year time period spanned the 1968 to 1997 water years, inclusive. The water year in Los Angeles is defined as April 1 to March 31. The results were analyzed to determine the average annual runoff volumes and TSS loads for each catchment. Figure 1 displays the average annual runoff volumes from the three catchments for the two LULC datasets. On average, the two datasets provide comparable annual runoff volume predictions with percent differences of 8, 14, and 9 for Centinela, Sepulveda, and Ballona, respectively. Figure 2 shows a comparison of the average annual TSS loads from the three catchments using the two different LULC datasets. Using the USGS dataset consistently provides a higher TSS prediction. The same trend was noticed for runoff volume, but the magnitude of the difference is greater for TSS.

Table 5. Mapping of LADPW monitored land uses to simulated land uses.

Simulated Land Use	LADPW Monitored Land Use
Commercial	Commercial
Residential	High-Density Single Family Residential
Transportation	Transportation
Industrial	Industrial
Urban Vacant	Vacant
Other Urban	Average of Commercial, Industrial, and Transportation EMCs
Water	Zero Concentration (does not contribute appreciably to runoff)

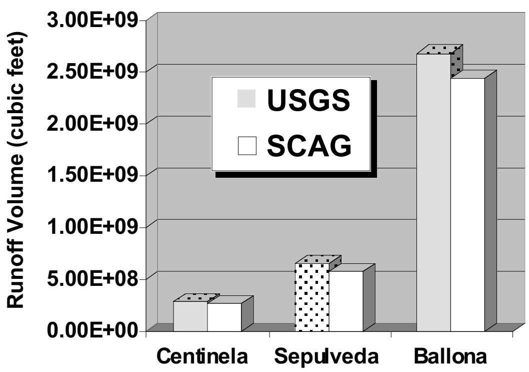

Figure 1. Comparison of average annual runoff volume predictions using USGS versus SCAG LULC datasets.

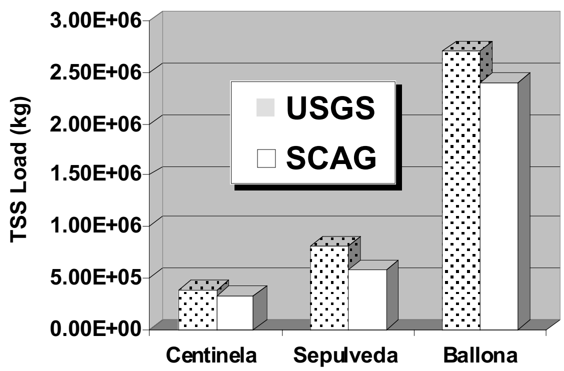

Figure 2. Comparison of average annual TSS load predictions using USGS versus SCAG LULC datasets.

SUMMARY

Most urban watershed models are dependent on land use data for simulating stormwater runoff quantity and quality. Using accurate land use datasets with the appropriate spatial resolution and level of detail are imperative to produce accurate simulation results. Traditionally the USGS LULC datasets have been used for most large-scale environmental modeling applications in the U.S. But, the USGS LULC datasets, in general, do not have the accuracy, level of detail, or spatial resolution necessary for many present day urban environmental modeling applications. For example, evaluation of land-use specific BMP implementation in alternative stormwater management plans requires high resolution, up-to-date land use information, which in most cases is beyond the limitations of the USGS dataset. In response to the need for improved datasets, local/regional governments, commercial entities, as well as the USGS have been active in developing newer, higher resolution datasets.

In this paper several potential sources of urban LULC data were briefly described including those from the USGS and a regional government. An investigation of the relative impact of using the USGS LULC dataset versus a more accurate regional government dataset for urban watershed modeling for three catchments in Los Angeles was performed. The results were summarized into average annual runoff volumes and TSS loads. When using the USGS LULC dataset compared to the regional government dataset the average annual runoff volume and TSS load were predicted to be 8% higher and 18% higher for the 24-km^2 Centinela catchment, 14% higher and 40% higher for the 61-km^2 Sepulveda catchment, and 9% higher and 13% higher for the 217-km^2 Ballona catchment. These results tend to suggest that even for urban areas that have not undergone significant land use change in the past two decades (e.g., Ballona Creek watershed) significant differences in runoff volume and pollutant load prediction can be expected when using the USGS LULC dataset versus other datasets.

REFERENCES

Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data*. USGS Professional Paper 964, U.S. Geological Survey.

Bannerman, R. T., Owens, D. W., Dodds, R. B. and Hornewer, N. J. (1993). Sources of pollutants in Wisconsin stormwater. *Wat. Sci. and Tech.*, **28**(3-5), 241-259.

Bhaduri, B., Harbor, J., Engel, B. and Grove, M. (2000). Assessing watershed-scale, long-term hydrologic impacts of land-use change using a GIS-NPS model. *Environmental Management*, **26**(6), 643-658.

Burian, S. J., McPherson, T. N., Brown, M. J. and Turin, H. J. (2000). *Development of a stormwater model for the Ballona Creek watershed*. LA-UR-00-1849, Los Alamos National Laboratory, Los Alamos, NM.

Charbeneau, R. J. and Barrett, M. E. (1998). Evaluation of methods for estimating stormwater pollutant loads. *Water Environment Research*, **70**(7), 1295-1302.

Dale, V. H., Brown, S., Haeuber, R. A., Hobbs, N. T., Huntly, N., Naiman, R. J., Riebsame, W. E., Turner, M. G. and Valone, T. J. (2000). Ecological principles and guidelines for managing the use of land. *Ecological Applications*, **10**(3), 639-670.

EPA (U.S. Environmental Protection Agency). (1983). *Results of the Nationwide Urban Runoff Program. Volume I. Final Report*. Water Planning Division, Washington, DC.

Harbor, J. (1994). A practical method for estimating the impact of land use change on surface runoff, groundwater recharge and wetland hydrology. *Journal of the American Planning Association*, **60**, 90-104.

Hill, K., Botsford, E. and Booth, D. (2000). *A rapid land cover classification method for use in urban watershed analysis*. Report downloaded from the Center for Urban Watershed Management Web Site (http://depts.washington.edu/cuwrm/).

Huber, W. C. and Dickinson, R. E. (1988). *Storm Water Management Model, Version 4, Part A: User s Manual*. EPA-600/3-88-001a, U.S. Environmental Protection Agency, Athens, Georgia.

LADPW (Los Angeles County Department of Public Works). (2000). *Integrated receiving water impacts report*.

Marsalek, J. (1978). *Pollution due to urban runoff: Unit loads and abatement measures*. Pollution from Land Use Activities Reference Group, International Joint Commission, Windsor, Ontario.

Novotny, V. and Olem, H. (1994). *Water quality: Prevention, identification, and management of diffuse pollution*. Van Nostrand Reinhold, New York, NY.

Schueler, T. R. (1987). *Controlling urban runoff: A practical design manual for planning and designing urban BMPs*. Metropolitan Washington Council of Governments.

USGS (U.S. Geological Survey). (1990). *Data users guides*. Department of the Interior, Reston, VA.